International Journal of Production Economics 206 (2018) 261–267

Contents lists available at ScienceDirect

International Journal of Production Economics

journal homepage: www.elsevier.com/locate/ijpe

Optimal sizing and planning of onsite generation system for manufacturing in Critical Peaking Pricing demand response program

Yunchao Zhang[a], Md Monirul Islam[b], Zeyi Sun[b,*], Sijia Yang[c,d], Cihan Dagli[b], Haoyi Xiong[a]

[a] Department of Computer Science, USA
[b] Department of Engineering Management and Systems Engineering, Missouri University of Science and Technology, USA
[c] Telecom Ecole de Management, France
[d] Beijing University of Posts and Telecommunications, China

ARTICLE INFO

Keywords:
Onsite generation system
Sizing
Manufacturing system
Genetic algorithm
Critical peak pricing
Demand response

ABSTRACT

Onsite electricity generation system in manufacturing has been traditionally considered an effective backup energy source to support the manufacturing operations when external power is not available due to natural disasters and/or power blackouts. Recently, with the increasing concerns of climate change and environmental protection, the contribution of using onsite generation system (OGS) to the manufacturing end use customers when they enroll in specific electricity demand response programs has also been gradually recognized. In this paper, we investigate the cost-effective OGS sizing problem for manufacturing practitioners when participating in Critical Peaking Pricing (CPP) demand response program. A Mixed Integer Non-Linear Programming (MINLP) formulation is proposed to identify the optimal size and utilization strategy of the OGS, as well as the corresponding production plan of the manufacturing system to minimize the overall energy related cost. Linearization strategy and metaheuristic algorithm are discussed for solving the proposed formulation with a reasonable computational cost and a good solution quality. A case study based on a real auto component manufacturing system and an existing CPP program is implemented to examine the effects of the proposed model. The results show that when utilizing the OGS appropriately sized, the total electricity related cost of the manufacturing system can be significantly reduced when participating in the CPP program.

1. Introduction

Onsite electricity generation system has been traditionally considered an effective backup energy source to improve the resilience of facility via supplying energy when external power is not available due to natural disasters and/or power blackouts. A great number of OGSs have been constructed for various critical infrastructure (New York State Division of Homeland Security and Emergency Services, 2014; Stadler, 2014) as well as residential housing (Ahourai and Faruque, 2013; Roggia et al., 2011; Hawkes and Leach, 2007).

Recently, the applications of OGS for manufacturers have also been launched to improve the resilience of the manufacturing systems against grid outage (Zhong et al., 2017; John, 2016). In addition, with the growing strong concerns of climate change, the application of using OGS by manufacturers in a carbon-constrained world has been gradually recognized, and thus, the research in this area has been launched (Xia et al., 2012). One typical application is to use the OGS as a critical alternative energy source for manufacturing plants when they participate in electricity demand response programs. Onsite generated energy can be provided to the manufacturing system in peak periods when demand response event occurs typically during the periods when the supply of the electricity cannot meet all the demand throughout the grid.

Electricity demand response is thought of as a critical strategy from the customer side to balance the electricity generation and consumption throughout the grid. It is defined as “changes in electric usage by end-use customers from their normal consumption patterns in response to the changes in the price of electricity over time, or to the incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized” (Federal Energy Regulatory Commission, 2012). The program urges various electricity consumers to alter their respective regular consumption patterns as a response to the changes of electricity price along the time horizon.

The studies focusing on the policy-making for promoting demand response programs have been reported (Greening, 2010; Vassileva

* Corresponding author.
E-mail address: sunze@mst.edu (Z. Sun).

https://doi.org/10.1016/j.ijpe.2018.10.011
Received 30 April 2018; Received in revised form 14 August 2018; Accepted 22 October 2018
Available online 24 October 2018
0925-5273/

et al., 2012). Also, the literature focusing on facilitating demand response for the stakeholders of generation and consumption sides of the grid can be found. For the generation side, the electricity price design tools (Doostizadeh and Ghasemi, 2012; Faria et al., 2011; Yousefi et al., 2011; Yu et al., 2012), as well as the simulation-based benefit evaluation mechanisms (Faria and Vale, 2011) have been studied. For the consumption side, a lot of research investigating the participation strategy of demand response for the customers in commercial and residential sectors has been conducted. For example, an optimal real time demand response decision-making model for residential house was studied in (Chen et al., 2012; Yi et al., 2013). An intelligent agent used in smart building for balancing the tradeoff between user comfort and energy consumption was proposed in (Corno and Razzak, 2012).

For the industrial end use customers, the relevant research has also been reported in recent years (Islam et al., 2017, 2018; Shrouf et al., 2013; Logenthiran et al., 2012). For instance, Islam et al. proposed a simulation-based demand response implementation strategy for manufacturers with OGS (Islam et al., 2017). An optimization model was proposed to identify the schedule of a single manufacturing machine to minimize the electricity billing cost (Shrouf et al., 2013).

Generally, there exist two different types of programs of demand response, i.e., price-based and event-based (Goldman et al., 2010). For the price-based program, the electricity rates fluctuate along the time horizon to urge the customers to change their original styles of energy consumption. It aims to balance the supply and demand from a long-term perspective. The most prevalently used price-based program is Time-of-use (TOU) tariff mechanism (Federal Energy Regulatory Commission, 2012). Unit electricity consumption rate varies depending on the time periods. Some TOU tariffs also include the demand charge based on the maximum power drawn from the grid throughout the billing cycle, typically, a month, in addition to electricity consumption charge.

For the event-based program, the customers are expected to be rewarded by reducing the power level drawn from the grid when requested by the utility due to specific triggering events, e.g., extreme high temperature in a summer afternoon, hardware equipment breakdowns, etc. It aims to reduce the load on a short-term basis upon the occurrence of such events. The occurrence frequency and the duration of the event vary depending on different applications (Goldman et al., 2010). Critical Peak Pricing (CPP) program is a typical event-based program. An extremely high rate of energy consumption is applied for the periods when “critical peaks” occur during which the electricity demand might be extremely high, while a discounted price for the time periods of the remaining time is also offered to the customers (Chino Valley Unified School D, 2012). A lot of pioneering customers pursuing sustainable operation mode are in favor of the CPP applications (Coastal Pacific Food Distributors, 2012; Railex, 2012; The Water Garden, 2010; California Public Utilities Commission, 2013).

In comparison to traditional TOU that has already been widely used in many areas of the United States as a base electricity tariff system, CPP program is comparatively new, but it has obtained more and more attention. Many utility companies, for example, Pacific Gas & Electric (PGE), Southern California Edison (SCE), and San Diego Gas & Electric (SDGE) (California Public Utilities Commission, 2014; Wang and Li, 2016), have started their implementation of CPP program by providing CPP as an optional electric service to complement the existing TOU systems. In a recent survey, Wang and Li analyzed CPP programs used by manufacturing applications in California. The CPP characteristics are summarized based on historical records (Chino Valley Unified School D, 2012). It indicated that for manufacturing end use customers with energy consumption adjustment flexibility by appropriately rescheduling their production, the average electric billing cost can be reduced by adopting CPP.

In this paper, we further extend the research of manufacturers' participation in CPP (Chino Valley Unified School D, 2012) by considering OGS. A recent paper employing simulation method illustrated the benefits of using OGS for manufacturers in demand response (Islam et al., 2017). However, the paper assumed an existing OGS with a given size and did not handle the details in the environment of CPP program. Thus, the objective of this paper is to investigate the economic viability of sizing an OGS for manufacturers when considering the participation of CPP program. Mixed Integer Non-Linear Programming (MINLP) is used to build the formulation so that the size and utilization strategy of the OGS for manufacturing as well as the corresponding optimal production plan minimizing the overall energy related cost when participating in CPP program can be optimally identified. The remaining part of this paper is organized as follows. The MINLP formulation of the proposed sizing model is given in Section 2. Different solution strategies are then discussed and compared in Section 3. After that, a case study based on a real manufacturing system and an existing CPP program is carried out in Section 4. Finally, the paper is concluded and the future work is discussed in Section 5.

2. Sizing model

The special events that trigger the CPP may usually happen in the summer period of a year, e.g., from June to October, in the north hemisphere. For the manufacturers who participate in the CPP program, the credit of unit consumption charge ($/kWh) will be applied to only on-peak periods, both on-peak and mid-peak periods, or all periods depending on the specific programs. An additional unit consumption charge is applied for the energy consumed during the event durations. In addition, the credit of unit demand charge ($/kW) is also given to the participating manufacturers at different types of the periods.

Consider a typical manufacturing system where J sequential manufacturing tasks are conducted as shown in Fig. 1. Each task can be conducted by a few parallel manufacturing stations while each station may include multiple manufacturing machines to perform different manufacturing functionalities. The rectangles in Fig. 1 denote the manufacturing stations, while the circles in Fig. 1 denote the buffer locations where the work-in-progress parts can be stored. Let j be the indexes of the manufacturing tasks ($j = 1, \ldots, J$) and buffer locations

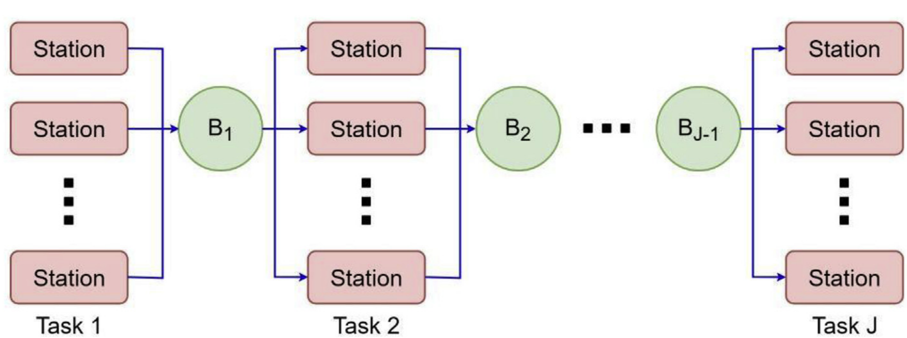

Fig. 1. A multi-task manufacturing system.

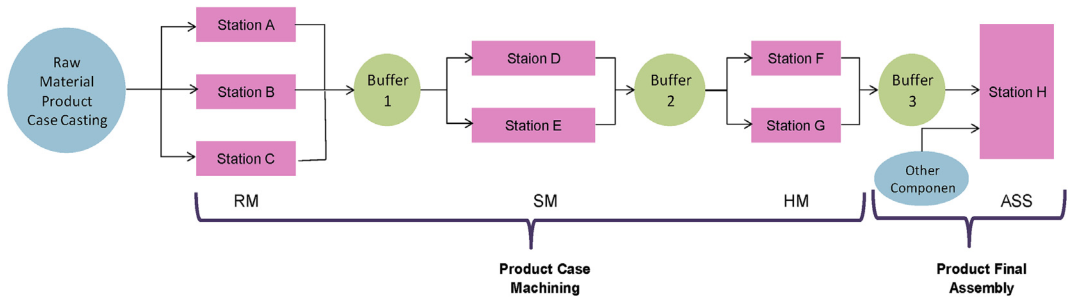

Fig. 2. Layout of an auto component manufacturing system.

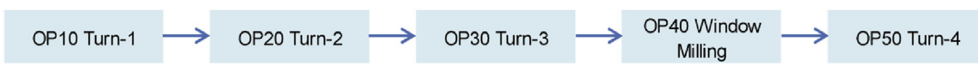

Fig. 3. An illustrative demonstration of Station A.

Table 1
Cycle time of stations.

Station	A	B	C	D	E	F	G	H
Cycle time (s)	135	135	135	80	80	80	80	40

Table 2
Buffer size and the content at the beginning.

	Capacity	Initial Content
Buffer 1	1800	400
Buffer 2	1800	400
Buffer 3	1800	400

Table 3
Rated power of the manufacturing machines in RMA.

Machine Name	Rated Power (kW)
OP10 Turn-1	21
OP20 Turn-2	21
OP30 Turn-3	21
OP40 Window milling	31
OP50 Turn-4	24

($j = 1, \ldots, J$-1). Let i be the index of the manufacturing stations in the manufacturing system. Let P_i be the rated power of station i in the manufacturing system.

Let m be the index of the months in a calendar year. The time horizon of each month includes a set of discretized intervals indexed by t_m ($t_m = 1, \ldots, T_m$) with constant duration of Δt. Also, we define z as the size of the OGS to be determined. We define x_{it_m} as the binary decision variable to represent the production plan of the manufacturing system. x_{it_m} is equal to one if the production plan specifies that the station i will produce in time interval t_m, while taking the value of zero if not. Let y_{t_m} be the decision variable denoting the energy offered by the OGS in time interval t_m. The purpose is to minimize the yearly energy relevant cost including both electricity billing cost and the onsite generation cost, which can be formulated by (1).

$$\min_{x_{it_m}, y_{t_m}, z} \sum_m (E_m + F_m) \tag{1}$$

where E_m and F_m are the electricity billing cost and the onsite generation cost in month m, respectively. E_m can be calculated by (2).

$$\begin{aligned} E_m = & \textstyle\sum_{t_m \in \mathbf{T_m} \setminus \mathbf{T_m^D}} \left[\left(\sum_i x_{it_m} P_i - y_{t_m}\right)\cdot \Delta t \cdot c_{t_m}^r\right] \\ & + \textstyle\sum_{t_m \in \mathbf{T_m^D}} \left[\left(\sum_i x_{it_m} P_i - y_{t_m}\right)\cdot \Delta t \cdot \left(c_{t_m}^r + c_{t_m}^a\right)\right] \\ & + (d_{on} - cr_{on})\cdot \max_{t_m \in \mathbf{T_m^{ON}}} \left(\textstyle\sum_i x_{it_m} P_i - y_{t_m}\right) + (d_{mid} - cr_{mid})\cdot \\ & \max_{t_m \in \mathbf{T_m^{MID}}} \left(\textstyle\sum_i x_{it_m} P_i - y_{t_m}\right) \\ & + (d_{off} - cr_{off})\cdot \max_{t_m \in \mathbf{T_m^{OFF}}} \left(\textstyle\sum_i x_{it_m} P_i - y_{t_m}\right) + (d_{base} - cr_{base})\cdot \\ & \max_{t_m \in \mathbf{T_m}} \left(\textstyle\sum_i x_{it_m} P_i - y_{t_m}\right) \end{aligned} \tag{2}$$

where $c_{t_m}^r$ is the electricity consumption charging rate ($/kWh) at interval t_m after the credit is applied due to the participation in CPP demand response program. $c_{t_m}^a$ is the additional charge rate for unit energy consumption during the intervals that belong to CPP demand response duration. d_{on}, d_{mid}, and d_{off} are the electricity demand charging rate ($/kW) for on-peak, mid-peak, and off-peak periods, respectively. d_{base} is the demand charging rate ($/kW) for all the time intervals throughout the billing cycle. Note that, in addition to the demand charge for peak, mid-peak, and off-peak periods, there also exists a base demand charge depending on the maximum power level drawn from the grid throughout all the time intervals in the month. cr_{on}, cr_{mid}, cr_{off}, and cr_{base} are the credits of electricity demand charging rates ($/kW) for on-peak, mid-peak, off-peak, and all the time intervals, respectively, when the manufacturer participates in the CPP demand response program. $\mathbf{T_m}$ is the set including all the time intervals in month m. $\mathbf{T_m^D}$ is the set including the time intervals when CPP demand response events happen in month m. $\mathbf{T_m^{ON}}$, $\mathbf{T_m^{MID}}$, and $\mathbf{T_m^{OFF}}$ are the sets including the time intervals belonging to on-peak, mid-peak, and off-peak periods, respectively, in month m.

On the right hand side of (2), the first two terms mean the electricity consumption charge in month m. Depending on the occurrence of CPP demand response events, two different unit consumption charging rates are applied. The remaining four terms are the electricity demand charge throughout the entire month.

F_m can be calculated by (3).

$$F_m = \sum_{t_m \in \mathbf{T_m}} f \cdot y_{t_m} \tag{3}$$

where f is the operation, maintenance, and fuel cost of the OGS for generating unit electricity.

The constraints of the problem are formulated as follows. The customer demand is represented by the purchase order from the customers with specified periodic shipment quantity, thus, the total production at each shipment cycle (i.e., between two required shipment time points), in practice, typically needs to be no less than the shipment quantity specified by the customer purchase order. It is formulated by (4).

$$\sum_{t \in \mathbf{T_k}} \sum_{i \in \mathbf{J}} x_{it} \cdot PR_i \cdot \Delta t \geq A_k \tag{4}$$

where $\mathbf{J}$ is the set of the manufacturing stations executing manufacturing task J. PR_i is the production rate of the station i. A_k is the

Table 4
Charging rates of SCE TOU-8.

Months	Energy Charge ($/kWh)			Demand Charge ($/kW)			Monthly Charge ($)
	On-peak	Mid-peak	Off-peak	On-peak	Mid-peak	Base (Any time)	
Jun–Sep	0.15267	0.09289	0.06592	25.16	7.11	14.99	596.11
Oct–May		0.09454	0.07165				

Table 5
Credits in SCE TOU-8 for participating customers.

Discount Months	Demand Charge Credits ($/kW)	Adder During CPP Events ($/kWh)	Event Time Period
Jun–Sep	On-Peak: 11.9	1.37	14:00–18:00

Table 6
Optimization results ($) comparison between GA and MILP.

	GA	MILP
Baseline	83,627.32	75,150.40
Proposed Method	48,411.6	37,964
Reduction (%)	42%	49%

Table 7
Computational time (s) comparison between GA and MILP solver.

	Baseline	Proposed Method
GA	29,600	35,016
MILP	125,140	200,920
Difference (%)	−76.3	−82.6

required shipment quantity specified by the customer purchase order at shipment cycle k. $\mathbf{T_k}$ is the set including the time intervals belong to the shipment cycle k. Note that the subscript m is removed from subscript t in (4), since the shipment does not have to be organized on a monthly basis. We also assume the shipment occurs at the first discretized time interval belongs to $\mathbf{T_k}$.

The energy output from the OGS needs to be no higher than the energy demand from the manufacturing system and the designed size of the OGS, whichever is smaller. It can be formulated by (5).

$$y_{t_m} \le \min\left(\sum_i x_{it_m} P_i,\ z\right),\ \forall\ t_m \tag{5}$$

The work-in-progress parts stored in the buffer locations should be maintained between the required minimum buffer stock and the capacities of the buffer locations, which can be formulated by (6).

Table 8
Cost reduction between the proposed model and baseline model.

Yearly Energy Cost ($)		Reduction (%)
Baseline	Proposed Method	
288,994.24	174,652.5	39.6

$$C_{j\_\min} \le B_{jt_m} \le C_{j\_\max} \tag{6}$$

where C_{j_min} is the minimum level of work-in-progress parts need to be maintained at buffer location j and C_{j_max} is the capacity of buffer location j. B_{jt_m} is the buffer content in buffer location j at the beginning of interval t_m, which can be recursively calculated by (7).

$$B_{jt_m} = B_{j(t_m-1)} + \sum_{i\in\mathbf{j}} x_{i(t_m-1)}\cdot PR_i\cdot\Delta t - \sum_{i\in\mathbf{j+1}} x_{i(t_m-1)}\cdot PR_i\cdot\Delta t,\ \forall\ t_m,\ \forall\ j \tag{7}$$

where $\mathbf{j}$ is the set of the manufacturing stations executing manufacturing task j.

The OGS typically has hardware operational constraints such as minimum ON (OFF) time after the system is activated (turned off), which can be described by (8).

$$\begin{cases} y_{t_m} > 0, & if \quad 1 \le V_{t_m} < V_{on} \\ y_{t_m} = 0, & if \quad -1 \ge V_{t_m} > -V_{off} \end{cases} \quad t_m = 2,\ \dots,\ T_m \tag{8}$$

In (8), V_{on} is the shortest required ON duration of the OGS after start-up; V_{off} is the shortest required OFF duration of the OGS after shutdown; and V_{t_m} is the consecutive ON or OFF duration of OGS up to the beginning of interval t_m. A positive (negative) value is used for the consecutive ON (OFF) time. V_{t_m} can be formulated in (9).

$$V_{t_m} = \begin{cases} \max(V_{t_m-1},\ 0) + 1, & \text{if} \quad y_{t_m-1} = 1 \\ \min(V_{t_m-1},\ 0) - 1, & \text{if} \quad y_{t_m-1} = 0 \end{cases} \quad t_m = 2,\ \dots,\ T_m \tag{9}$$

The initial condition of V_{t_m} is assumed to be $V_1 = 0$.

The inventory of the finished products needs to be maintained within a given range at the end of each interval t_m, which can be formulated by (10).

$$L \le Q_{t_m} \le U \tag{10}$$

where Q_{t_m} is the inventory of the finished products at the end of interval

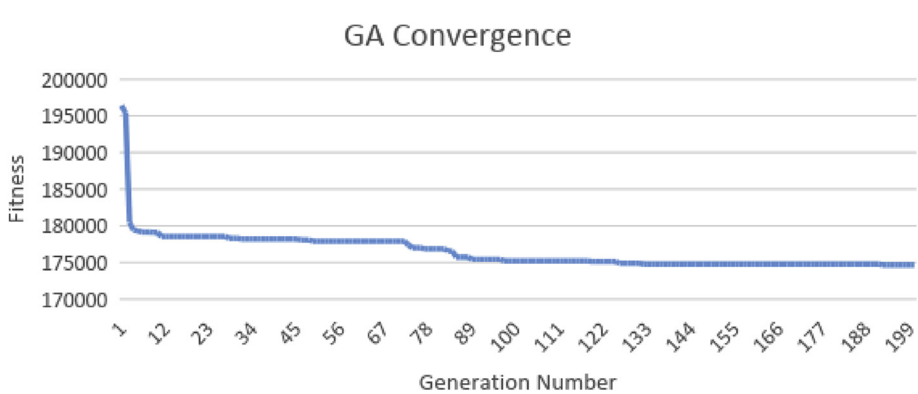

Fig. 4. GA convergence of the proposed model.

Table 9
Sensitivity analysis.

Parameters	Discount Rate (%)			Lifetime (years)		
Level	1	3	5	15	20	25
Cost Reduction (%)	39.60	39.56	39.76	39.70	39.56	39.78
Parameters	Demand charge credit ($/kW)			Adder during CPP ($/kWh)		
Level	10.9	11.9	12.9	1.07	1.37	1.67
Cost Reduction (%)	39.50	39.56	39.30	38.74	39.56	40.87

t_m after shipping the final products to satisfy the customer demand if required. L is the lower bound, while U is the upper bound of the inventory of finished products. Q_{t_m} can be recursively calculated by (11).

$$Q_{t_m} = Q_{t_{m-1}} + \sum_{i\in \mathbf{J}} x_{it_m}\cdot PR_i\cdot\Delta t - S_{t_m} \tag{11}$$

where S_{t_m} is the shipment quantity of the final products at the beginning of interval t_m. S_{t_m} can be calculated by (12).

$$S_{t_m} = \begin{cases} A_k, & \text{if shipment } A_k \text{ is required at interval } t_m \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

The net present value (NPV) of building an OGS with a size of z over the expected lifetime needs to be positive. It can be represented by (13).

$$NPV = -z\cdot H + \sum_{g=1}^{G} \frac{SA}{(1+r)^g} > 0 \tag{13}$$

where H is the setup cost of the OGS per unit size. r is the yearly discount rate. G is the lifetime of the OGS. g is the index of the year. SA is the yearly saving due to the use of OGS, which can be calculated by (14).

$$SA = B' - \left(\sum_m E_m + F_m\right) \tag{14}$$

where B' is the yearly electricity billing cost without OGS. It can be obtained through running the revised optimization model by removing OGS related decision variables, parameters, objectives, and constraints, i.e., to identify an optimal production plan that can minimize the yearly electricity billing cost when participating in the CPP program. Here we assume the model parameters across the years within the lifetime are the same and thus the yearly saving is constant.

3. Solution strategy

It can be seen that the sizing model proposed in Section 2 is a Mixed Integer Non-Linear Programming (MINLP). To solve this MINLP model, we can either employ a certain software solver that is typically applicable to the linearity formulation to obtain the optimal solution or use a certain meta-heuristic algorithm to obtain a near optimal solution considering the balance between solution qualities and computing cost. Various linearization strategies have been proposed to linearize the MINLP formulation to Mixed Integer Linear Programming (MILP) formulation (Sirvent et al., 2017; Hamzeei and Luedtke, 2014; Belotti et al., 2012) so that the existing solver requiring the assumptions of "differentiable", "convexity", and "continuous" can be used. While, the meta-heuristic algorithms can be used directly without such assumptions. For example, the evolutionary-based meta-heuristics, like Genetic Algorithm (Eiben et al., 1994) and Particle Swarm Optimization (Kennedy et al., 2001) have been extensively studied and widely used when solving some high-dimension complex optimization problems (Jerald et al., 2005; Du et al., 2015; McCall, 2005; Liu et al., 2014). Some researchers have recently explored the theoretical connection between such heuristic algorithms and some existing analytical models such as Hamiltonian systems and Nesterov's method (Freidlin and Hu, 2011; Hu and Li, 2017). In this section, we investigate and discuss the solution strategies considering both pathways aforementioned.

3.1. Linearization

We first adopt the linearization strategies as follows to linearize the non-linear terms in (2), (5), and (8), so that some existing linear programming solvers can be used to identify the optimal solutions.

For the "max" operator in (2), it is linearized by (15).

$$\begin{aligned} E_m = & \textstyle\sum_{t_m\in \mathbf{T_m}\setminus\mathbf{T_m^D}} [(\sum_i x_{it_m}P_i - y_{t_m})\cdot\Delta t\cdot \\ & c^r_{t_m}] + \textstyle\sum_{t_m\in\mathbf{T_m^D}} [(\sum_i x_{it_m}P_i - y_{t_m})\cdot\Delta t\cdot(c^r_{t_m} + c^a_{t_m})] \\ & + (d_{on} - cr_{on})\cdot h_{ON} + (d_{mid} - cr_{mid})\cdot h_{MID} + (d_{off} - cr_{off})\cdot h_{OFF} \\ & + (d_{base} - cr_{base})\cdot h_{BASE} \end{aligned} \tag{15}$$

In (15), the "max" operator is substituted by h_{ON}, h_{MID}, h_{OFF}, and h_{BASE}. Given the new variables, some additional constraints need to be formulated as follows.

$$\begin{aligned} & h_{ON} \geq \textstyle\sum_i x_{it_m}P_i - y_{t_m},\ \ \forall\ t_m \in \mathbf{T_m^{ON}} \\ & h_{MID} \geq \textstyle\sum_i x_{it_m}P_i - y_{t_m},\ \ \forall\ t_m \in \mathbf{T_m^{MID}} \\ & h_{OFF} \geq \textstyle\sum_i x_{it_m}P_i - y_{t_m},\ \ \forall\ t_m \in \mathbf{T_m^{OFF}} \\ & h_{BASE} \geq \textstyle\sum_i x_{it_m}P_i - y_{t_m},\ \ \forall\ t_m \in \mathbf{T_m^{BASE}} \end{aligned} \tag{16}$$

For the purpose of linearizing (5), we disaggregate it to two linear constraints

$$y_{t_m} \leq \sum_i x_{it_m}P_i,\ \ \forall\ t_m \tag{17}$$

$$z \geq y_{t_m},\ \ \forall\ t_m \tag{18}$$

For the nonlinearity in (8), (19) is used for linearization. In (19), two auxiliary binary variables, i.e., $l_{on}^{t_m}$ and $l_{off}^{t_m}$, are defined to denote if the ON/OFF status of the OGS is changed or not at the beginning of interval t_m. $l_{on}^{t_m}$ and $l_{off}^{t_m}$ take the value of one if the system starts and shuts off, respectively, at the beginning of interval t_m, and zero otherwise.

$$\begin{aligned} & l_{on}^{t_m} \leq y_{t_{m+q}},\ \ t_m = 2,\dots,T_m,\ \ \forall\ q\in\{1..V_{on}-1\} \\ & l_{off}^{t_m} \leq 1 - y_{t_{m+q}},\ \ t_m = 2,\dots,T_m,\ \ \forall\ q\in\{1..V_{off}-1\} \end{aligned} \tag{19}$$

After applying the linearization strategy aforementioned, the MINLP problem can be transformed to a Mixed Integer Linear Programming (MILP) problem. The toolbox "OR-Tools" from Google Optimization Tools is used to solve the MILP. The algorithm used in the toolbox implements Coin or Branch and Cut (CBC) as the solver. OR-Tools can work very effectively to solve the proposed problem with a three-month (e.g., June, July, and August, since most CPP events occur in the months of June, July, and August in the United States) time scale using 12,514 s.

However, when the time scale is increased from three months to 1 year as intended by the proposed problem, due to the computational complexity of the problem, the time required by the solution increased dramatically and the toolbox becomes inefficient. The number of variables is increased from 6,261,197 to 98,710,310. Further, the

amount of memory required for linear optimization also increases by a magnitude. Thus, the toolbox becomes impractical due to extremely long computational time. Although we tried to use a high performance cluster computer of dual 16 core Haswell CPUs with 256 GB DDR4 RAM for the computation, which took over 168 h without obtaining the solution.

3.2. Implementing the Genetic Algorithm

Due to the high dimension of the proposed model and the inefficiency of using linear solver toolbox to solve the problem after linearization, we propose to use a widely used meta-heuristic method, Genetic Algorithm (GA), to solve the proposed problem to obtain a near optimal solution with a reasonable computational cost. GA is inspired by the process of natural selection that belongs to the larger class of evolutionary algorithms, which has been widely used in high dimension scheduling problem involving binary decision variables.

We discretize the optimization horizon, i.e., one year, into $\sum_m T_m$ same duration intervals during which the production plan of the manufacturing system with D manufacturing stations and the utilization strategy of the OGS needed to be encoded. Thus, here we represent the production plan of each station and the utilization strategy of OGS as one chromosome with $\sum_m T_m$ genes. Together, they form a genotype of $(D+1)\sum_m T_m$ genes. The genes in the D chromosomes for production plan are encoded as binary variables (0, 1) and the genes in the chromosome for the OGS are encoded as double values within the range between zero and the maximum demand of the manufacturing system.

Considering that the search space of this problem is fairly large, we initially generate a population with a certain number of genotypes randomly. Then, we use tournament selection method to select the fitter parent genotypes while keeping the diversity of the population. For evolution process, uniform crossover and mutation are introduced for a broader exploration of the search space. The evolution is allowed to continue for a given number of generations. The fitness of each genotype can be calculated by (20) where all the constraints are integrated as penalty terms. After the last round of evolution, the genotype that shows the best fitness is returned as the result.

$$\begin{aligned}&\sum_m (E_m+F_m) + W\Bigg(\sum_k \left[\min\left(\sum_{t\in \mathbf{T_{k-1}}}\sum_{i\in \mathbf{J}} x_{it}\cdot PR_i\cdot\Delta t - A_k,\ 0\right)\right]^2 \\ &+ \sum_{t_m}\max(y_{t_m} - \min(\textstyle\sum_i x_{it_m}P_i,\ z),\ 0) \\ &\quad+ \sum_j\sum_{t_m}[\min(B_{jt_m} - C_{j\_\min},\ 0)]^2 \\ &+ \sum_j\sum_{t_m}\max(B_{jt_m} - C_{j\_\max},\ 0) + \sum_{t_m}\max(Q_{t_m} - U,\ 0) \\ &+ \sum_{t_m}[\min(Q_{t_m} - L,\ 0)]^2 + [\min(NPV,\ 0)]^2 \\ &+ \sum_{t_m=2}^{T_m}[\min(|y_{t_m} - y_{t_m-1}|\cdot|V_{t_m-1}| - y_{t_m-1}\cdot V_{on} - (1-y_{t_m-1})\cdot V_{off},\ 0)]^2\Bigg)\end{aligned} \tag{20}$$

In (20), W is a large positive real number to scale up the penalty due to the violations of the constraints.

4. Case study

The manufacturing system we use in this case study is an auto component manufacturing system with the layout as shown in Fig. 2.

Two major processes, i.e., machining and assembly are included in this auto component manufacturing system. The machining process can be further disaggregated to three sequential sub-processes (i.e., RM, SM, and HM). First, the initial surface processing on the raw materials of castings is fulfilled by RM. Next, the casting surface cutting and drilling are fulfilled by SM. Finally, the final finishing of the castings is completed by HM. In RM, Stations A/B/C are deployed in parallel, while in SM and HM, two machining stations, i.e. Station D/E, and Station F/G are deployed in parallel, respectively. Various computer numerical controlled (CNC) machines with respective purposes, e.g., grinding, drilling, turning, milling, etc., as well as several special machines with auxiliary purposes like balancing, demagnetization, and cleaning, are included in specific machining stations. An illustrative demonstration of Station A is show in Fig. 3.

The assembly process consists of a single assembly station denoted Station H to conduct assembly tasks. The assembly team conducts the required tasks in several workplaces deployed in assembly station to assemble the parts after machining along with the other components to finally complete the entire production process. The productivity related data (cycle time of each station, buffer capacity, buffer initial contents) is given in Tables 1 and 2. The rated power of the manufacturing machines in Station A is shown in Table 3 for illustration (the relevant data of other machines is not given due to confidentiality requirement).

The customer demand that needs to be satisfied on a weekly basis is 4800 units per week. The safety stock of the finished products requested by the customer needs to be maintained at the level that can cover one or two weeks' demand, i.e., between 4800 and 9600. The manufacturing system runs a five-day per week and two 8-h shifts per day working schedule. The production schedule and OGS planning are generated on an hourly basis. The minimum ON/OFF time of the OGS is 2 h. The lifetime of the OGS is 20 years. The yearly discount rate is set to be 3%.

Since the total rated power of the manufacturing system described in Fig. 2 is 720.5 kW, we select tariff "SCE TOU-8" (Southern California Edison, 2015), which is applicable for the customers with peak demand ranging from 500 kW to infinity, as the CPP demand response program in the case study. In "SCE TOU-8", there exists a fixed amount of 12 event days in one year and each event lasts 4 h. The charging rates at different periods and the credits for the participating customers are illustrated in Tables 4 and 5, respectively. We also retrieve the historical demand response events data of CPP in 2017 from SCE. Note that, due to the fact that CPP is relatively new, the existing historical dataset does not have enough size so that the proposed model cannot be run with stochastic inputs in term of various occurrence times of the CPP to obtain statistical results. Thus, we select the data from the most recent year in this case.

For comparison, we first run a baseline model without the OGS to examine the total energy billing cost under the same CPP demand response program. Then, we compare the optimization results and computing time between MILP after linearization and GA for a three-month case (June, July, and August) in Tables 6 and 7.

As we can see, although the cost reduction achieved by MILP is 7% better than GA as shown in Table 6, the computational time of MILP is much higher than GA. It implies that the proposed GA algorithm can obtain a near optimal solution by striking a balance between optimization quality and computational time.

Thus, we use GA to solve our proposed problem over one-year time scale. We generate an initial population of 50,000 genotypes and run the evolution algorithm for 500 generations. The GA convergence of the fitness is demonstrated in Fig. 4.

The size of the OGS is optimized to be 700 kW. The comparison of the yearly energy related costs between the proposed model with the OGS and the baseline model without the OGS is illustrated in Table 8. It shows that with an appropriate sized OGS, the electricity billing cost can be further significantly reduced under the CPP demand response program.

Further, sensitivity analysis is also implemented to examine the variations of the yearly energy cost reduction led by the change of modeling parameters including yearly discount rate, system lifetime, demand charge credit, and consumption charge adder during CPP as shown in Table 9. It can be seen that the cost reduction is fairly constant due to the variations of discount rate, lifetime, and demand charge credit. However, there is a slight increase of the saving percentage due to the increase of the additional consumption charge ($/kWh) during

CPP event duration. It indicates that the higher the penalty imposed on the energy consumption during CPP period by the program, the more cost reduction can be expected using the onsite generation system.

5. Conclusion and future work

In this paper, we investigate the economic sizing problem for the OGS used by manufacturers, especially for the participation in CPP demand response program. An MINLP model is proposed to identify the optimal OGS size and utilization strategy as well as the optimal production plan for the manufacturing system that can minimize the total energy related cost when manufacturing end use customers enrolled in CPP program. A real auto component manufacturing system and an existing CPP program are used in the numerical case study to examine and analyze the model effects. The case study shows that the installation and use of the OGS can further facilitate and promote the CPP program participation from manufacturer end use customers.

For future work, the sizing problem integrating the renewable sources like solar PV and wind turbine can be studied. Different demand response programs can also be examined. The influence due to the variations in customer demand in future years and the possible change of the energy cost parameters can be further tested.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpe.2018.10.011.

References

Ahourai, F., Faruque, M.A.A., 2013. Grid Impact Analysis of a Residential Microgrid under Various EV Penetration Rates in GridLAB-d. Technical report for Center for Embedded Computer Systems. University of California, Irvine.
Belotti, P., Kirches, C., Leyffer, S., Linderoth, J., Luedtke, J., Mahajan, A., 2012. Mixed-integer Nonlinear Optimization, Mathematics and Computer Science Division Preprint. Argonne National Laboratory ANL/MCS-P3060-1121 November 22, 2012.
California Public Utilities Commission, 2013. Commission Staff Report: Lessons Learned from Summer 2012 Southern California Investor Owned Utilities' Demand Response Programs. Retrieved from. http://www.cpuc.ca.gov/NR/rdonlyres/523B9D94-ABC4-4AF6-AA09-DD9ED8C81AAD/0/StaffReport_2012DRLessonsLearned.pdf.
California Public Utilities Commission, 2014. Demand Response. Retrieved from. http://www.cpuc.ca.gov/PUC/energy/Demand + Response/.
Chen, Z., Wu, L., Fu, Y., 2012. Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization. IEEE Trans. Smart Grid 3 (4), 1822–1831.
Chino Valley Unified School D, 2012. Energy Management Success Story: Chino Valley Unified School District. Retrieved from. https://www.sce.com/wps/wcm/connect/f6e4ba65-f019-4f9b-849c-4495112d4b9d/CSChinoValley_AA.pdf?MOD = AJPERES.
Coastal Pacific Food Distributors, 2012. Energy Management Success Story: Coastal Pacific Food Distributors. Retrieved from. https://www.sce.com/wps/wcm/connect/bc0210e9-60e4-48d3-8a27-51fdb2f50ce8/CSCoastalPacific_AA.pdf?MOD = AJPERE.
Corno, F., Razzak, F., 2012. Intelligent energy optimization for user intelligible goals in smart home environments. IEEE Trans. Smart Grid 3 (4), 2128–2135.
Doostizadeh, M., Ghasemi, H., 2012. A day-ahead electricity pricing model based on smart metering and demand-side management. Energy 46, 221–230.
Du, S., Huang, D., Wang, H., 2015. An adaptive support vector machine-based workpiece surface classification system using high-definition metrology. IEEE Trans. Inst. Meas. 64 (10), 2590–2604.
Eiben, A.E., et al., 1994. Genetic algorithms with multi-parent recombination. In: PPSN III: Proceedings of the International Conference on Evolutionary Computation. The Third Conference on Parallel Problem Solving from Nature, pp. 78–87 ISBN 3-540-58484-6.
Faria, P., Vale, Z., 2011. Demand response in electrical energy supply: an optimal real time pricing approach. Energy 36 (8), 5374–5384.
Faria, P., Vale, Z., Soares, J., Ferreira, J., 2011. Demand response management in power systems using a particle swarm optimization approach. IEEE Intell. Syst. 28 (4), 43–51.
Federal Energy Regulatory Commission, 2012. Reports on Demand Response & Advanced Metering, Washington D.C., U.S. Retrieved from. http://www.ferc.gov/industries/electric/indus-act/demand-response/dem-res-advetering.asp.
Freidlin, M., Hu, W., 2011. On perturbations of generalized Landau-Lifshitz Dynamics. J. Stat. Phys. 144, 978.
Goldman, C., Reid, M., Levy, R., Silverstein, A., 2010. National Action Plan for Energy Efficiency, Coordination of Energy Efficiency and Demand Response. United States Environmental Protection Agency, Washington D.C., U.S Retrieved from. http://www.epa.gov/cleanenergy/documents/suca/ee_and_dr.pdf.
Greening, L.A., 2010. Demand response resources: who is responsible for implementation in a deregulated market? Energy 35 (4), 1518–1525.
Hamzeei, M., Luedtke, J., 2014. Linearization-based algorithms for mixed-integer nonlinear programs with convex continuous relaxation. J. Global Optim. 59 (2–3). https://doi.org/10.1007/s10898-014-0172-4.
Hawkes, A.D., Leach, M.A., 2007. Cost-effective operating strategy for residential micro-combined heat and power. Energy 32, 711–723.
Hu, W., Li, J., 2017. On the Fast Convergence of Random Perturbations of the Gradient Flow. arXiv:1706.00837, Retrieved from. https://arxiv.org/abs/1706.00837.
Islam, M., Sun, Z., Dagli, C., 2017. Simulation-based investigations on electricity demand response for manufacturing systems to mitigate overgeneration due to high penetration of renewable source. DEStech Trans. Eng. Technol. Res. 226–231.
Islam, M., Zhong, X., Xiong, H., Sun, Z., 2018. Optimal scheduling of manufacturing and onsite generation systems in over-generation mitigation oriented electricity demand response program. Comput. Ind. Eng. 115, 381–388.
Jerald, J., Asokan, P., Prabaharan, G., Saravanan, R., 2005. Scheduling optimisation of flexible manufacturing systems using particle swarm optimisation algorithm. Int. J. Adv. Manuf. Technol. 25 (9), 964–971.
John, P., 2016. Design of solar energy system in food manufacturing environment, Mo. Cogent Eng. 3, 1233613. Retrieved from. https://doi.org/10.1080/23311916.2016.1233613.
Kennedy, J., Eberhart, R.C., Shi, Y., 2001. Swarm Intelligence. Morgan Kaufmann, CA.
Liu, C.Y., Zou, C.M., Wu, P., 2014. A task scheduling algorithm based on genetic algorithm and ant colony optimization in cloud computing. In: Proceedings of 2014 13th International Symposium on Distributed Computing and Applications to Business, Engineering and Science.
Logenthiran, T., Srinivasan, D., Shun, T.Z., 2012. Demand side management in smart grid using heuristic optimization. IEEE Trans. Smart Grid 3 (3), 1244–1252.
McCall, J., 2005. Genetic algorithms for modelling and optimization. J. Comput. Appl. Math. 184 (1), 205–222.
New York State Division of Homeland Security and Emergency Services, 2014. Microgrids for Critical Facility Resiliency in New York State. Retrieved from. https://www.nyserda.ny.gov/-/media/Microgrids-Report-Summary.pdf.
Railex, 2012. Energy Management Success Story: Railex. Retrieved from. https://www.sce.com/wps/wcm/connect/109348ba-7680-43f6-839f 6b7651b583ff/CaseStudy_Railex_AA.pdf?MOD = AJPERES.
Roggia, L., Rech, C., Schuch, L., Baggio, J.E., Hey, H.L., Pinheiro, J.R., 2011. Design of a sustainable residential microgrid system including PHEV and energy storage device. In: Proceedings of the 14th European Conference on Power Electronics and Applications (EPE 2011).
Shrouf, F., Ordieres-Meré, J., García-Sánchez, A., Ortega-Mier, M., 2013. Optimizing the production scheduling of a single machine to minimize total energy consumption costs. J. Clean. Prod. 67, 197–207.
Sirvent, M., Kanelakis, N., Geißler, B., Biskas, P., 2017. Linearized model for optimization of coupled electricity and natural gas systems. J. Mod. Power Syst. Clean Energy 5 (3), 364–374.
Southern California Edison, 2015. Rates & Pricing Choices. Retrieved from. www.sce.com/wps/portal/home/regulatory/tariff-books/rates-pricing-choices.
Stadler, M., 2014. Microgrid Planning and Operations for Critical Facilities Considering Outages Due to Natural Disasters. Retrieved from. https://building-microgrid.lbl.gov/sites/all/files/DER-CAM%20-%20microgrid%20resilience_V7.pdf.
The Water Garden, 2010. Energy Management Success Story: the Water Garden. Retrieved from. https://www.sce.com/wps/wcm/connect/0a346efc-9d6c-4b1a-8bf5-caf445b52a5a/WaterGardenCaseStudy.pdf?MOD = AJPERES.
Vassileva, I., Wallin, F., Dahlquist, E., 2012. Understanding energy consumption behavior for future demand response strategy development. Energy 46, 94–100.
Wang, Y., Li, L., 2016. Critical peak electricity pricing for sustainable manufacturing: modeling and case studies. Appl. Energy 175, 40–53.
Xia, T., Xi, L., Zhou, X., Du, S., 2012. Modeling and optimizing maintenance schedule for energy systems subject to degradation. Comput. Ind. Eng. 63 (3), 607–614.
Yi, P., Dong, X., Iwayemi, A., Zhou, C., Li, S., 2013. Real-time opportunistic scheduling for residential demand response. IEEE Trans. Smart Grid 4 (1), 227–234.
Yousefi, S., Moghaddam, M.P., Majd, V.J., 2011. Optimal real time pricing in an agent-based retail market using a comprehensive demand response model. Energy 36 (9), 5716–5727.
Yu, R., Yang, W., Rahardja, S., 2012. A statistical demand-price model with its application in optimal real-time price. IEEE Trans. Smart Grid 3 (4), 1734–1742.
Zhong, X., Islam, M., Xiong, H., Sun, Z., 2017. Design the capacity of onsite generation system with renewable sources for manufacturing plant. Proc. Comput. Sci. 114, 433–440.